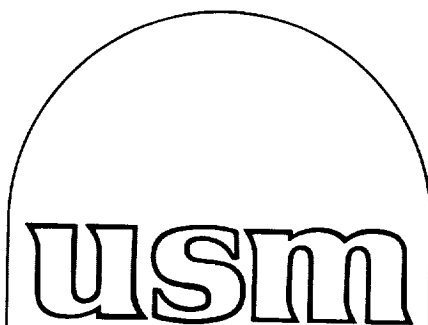


College of Science and Technology

11/16/86
66P



(NASA-CR-195748) CONTINUATION OF
THE DEVELOPMENT OF SOFTWARE TO BE
USED IN SUPPORT OF THE NASA
TECHNOLOGY UTILIZATION PROGRAM
(University of Southern
Mississippi) 66 p

N94-71790

Unclas

Z9/61 0002876

SCIENCE AND TECHNOLOGY

CONTINUATION OF THE DEVELOPMENT OF
SOFTWARE TO BE USED IN SUPPORT
OF THE NASA TECHNOLOGY UTILIZATION PROGRAM

July, 1986



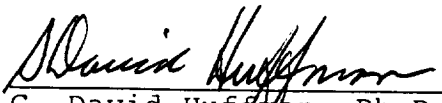
** rec'd 9/10/86*

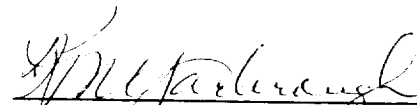
CONTINUATION OF THE DEVELOPMENT OF
SOFTWARE TO BE USED IN SUPPORT
OF THE NASA TECHNOLOGY UTILIZATION PROGRAM

G. David Huffman

University of Southern Mississippi
Hattiesburg, Mississippi 39406-5165

July, 1986


G. David Huffman, Ph.D.
Dean, College of
Science and Technology


K. Yarbrough, Ph.D.
Vice President for
Research and Extended Services

RESTRICTION ON USE AND DISCLOSURE OF
PROPOSAL AND QUOTATION INFORMATION

The information contained in pages 1-36 of this proposal constitutes a trade secret and/or information that is commercial or financial and confidential or privileged. It is furnished to the Government in confidence with the understanding that it will not, without permission of the offeror, be used or disclosed for other than evaluation purposes; provided, however, that in the event a contract is awarded on the basis of this proposal the Government shall have the right to use and disclose this information to the extent provided in the contract. This restriction does not limit the Government's right to use or disclose this information if obtained from another source without restriction. ||

EXECUTIVE SUMMARY

Technology transfer activities are carried out by NASA via a network of Industrial Applications Centers and State Technology Assistance Centers. These organizations employ a number of modes of technology transfer which principally vary in degree of specificity. These modes are largely the same regardless of the technology under consideration. The processes are labor intensive and are now being automated using micro-computer-based techniques and software support systems. The systems are currently being installed with testing, optimization and modification proposed for the coming contract year. Full implementation of the software support systems can reduce technology transfer study costs by up to 20-25%.

TABLE OF CONTENTS

<u>Section Number</u>	<u>Section Title</u>	<u>Page Number</u>
	Executive Summary	i
	Table of Contents	ii
	List of Figures	iii
	List of Tables	iii
	Nomenclature	iv
1	UNDERSTANDING OF THE PROBLEM	1
1.1	Bibliographic Database Searching and Technology Transfer	1
1.2	Survey of Commercial Software Systems	2
1.3	Software Framework for Technology Transfer	9
2	TECHNICAL APPROACH	11
2.1	Previous Work	11
2.1.1	User Facility and Common Command Language	11
2.1.2	Reviewing and Determining Relevancy of Abstracts	11
2.1.3	Translation of Citation Formats	12
2.2	SHORT-AID Post Processor	12
2.2.1	Program NABST	13
2.2.2	Program REVIEW	14
2.2.3	Program RANK	17
2.2.4	Program PRINT	17
2.3	Abstract Relevance Determination	18
2.4	Theory of Relevance Determination	18
2.5	Evaluation of Relevance Determination	22
2.6	Proposed Continuation of Work	22
2.7	Scope of Work	32
2.7.1	Work Statement	32
2.7.2	Schedule	33
2.7.3	Budget	33
2.7.4	Summary	36
3	PERSONNEL	37
3.1	Overview	37
3.2	Resume for G. David Huffman	38
3.3	Resume for G. Oehms	40
3.4	Resume for D. Vital	42
3.5	Resume for Deborah A. Theisen	43
4	OPERATING ORGANIZATION	44
4.1	The University of Southern Mississippi	44
4.2	The College of Science and Technology	48
4.3	The Department of Computer Science and Statistics	49
4.4	College Computer Facilities	49
4.5	Additional Equipment Required	50
5	EXPERIENCE AND PAST PERFORMANCE	52
6	REFERENCES AND BIBLIOGRAPHY	56

LIST OF FIGURES

<u>Figure Number</u>	<u>Figure Caption</u>	<u>Page Number</u>
1	Profile of Steps in an Industrial Applications Study	4
2	Typical Commercial Software Packages for Use in Technology Transfer	5
3	System Flow of SORT-AID	15
4	Occurrence Matrix	20
5	Keywords for Abstract Relevance, $w_k^{(f)}$	24
6	Keywords for Abstract Relevance, $w_k^{(s)}$	25
7	Keywords for Abstract Relevance, $w_k^{(b)}$	26
8	Abstract Distribution by Groups	27
9	Abstract Distribution by Groups	28
10	Program Schedule	34

LIST OF TABLES

<u>Table Number</u>	<u>Table Title</u>	<u>Page Number</u>
1	Descriptions of some Software Systems Relevant to Technology Transfer Tasks	6
2	Program Budget	35
3	Previous Software Development Contracts	53

NOMENCLATURE

<u>Variable</u>	<u>Description</u>
B_k	Number of documents in which k-th term occurs in J abstracts
$b_{j,k}$	Binary occurrence of k-th term in the j-th abstract
D_k	Discrimination factor for the k-th term
F_k	Number of occurrences of the term k in a collection of J abstracts
N_k	Noise of k-th term in J abstracts
$n_{j,k}$	Number of occurrences of the k-th term in the j-th abstract
R_j	Ranking of the j-th abstract
S_k	Signal of k-th term in J abstracts
$s_{i,j}$	Similarity measure between i-th and j-th abstract
V_k	Variance of k-th term
W_k	Weighting factor for the k-th term
$\beta_{j,k}$	Relevance measure for k-th term in j-th abstract
λ_k	Relevance of k-th term

SUBSCRIPTS

<u>Variable</u>	<u>Description</u>
i	Abstract
J	Total number of abstracts
j	Abstract
K	Total number of terms
k	Term

SUPERSCRIPTS

<u>Variable</u>	<u>Description</u>
b	Binary frequency
f	Frequency
s	Signal
v	Variance
β	Similarity

1. UNDERSTANDING OF THE PROBLEM

1.1 Bibliographic Database Searching and Technology Transfer

The NASA technology transfer process is regarded as an information system. A primary output is known as the "industrial applications study". The actual study processes involved in the industrial applications study are shown in flow chart format in Figure 1. The first step in the process is the formulation of the problem statement. Once the problem is formulated, a series of databases are selected. These may be either commercial, e.g., DIALOG, or public domain, e.g., NASA RECON. Using database thesauri, the search strategy is formulated and a series of keywords chosen. An on-line search is conducted and a number of abstracts retrieved. These are then reviewed for relevance and a number of documents ordered. The documents can then be analyzed and, if warranted, government and/or industrial contacts instituted. Information from experts in conjunction with the assessment of the published information is used to formulate the final report.

Another output of the NASA process is the "current awareness search". The steps carried out in a current awareness search are similar, but are normally terminated after the on-line search, i.e., the final report consists of the retrieved abstracts. Presumably, the recipient of the abstracts carries out the abstract review and analysis procedure and contacts the relevant experts.

The labor intensive steps in the above process are associated with the review of the abstracts for relevance, the

analysis of the relevant publications and the preparation of the final report. Analysis and report writing labor can be reduced through the employment of computer information systems tools and techniques. This is discussed in the following sections.

1.2 Survey of Commercial Software Systems

Software development programs which are relevant to technology transfer are being carried out by government laboratories, businesses and universities. The various programs are diverse and difficult to categorize. They range from highly focused products for IBM PC's such as IN-SEARCH, e.g., Menlo (1984), to a broad-based program such as the Total Information System (TIS)--a DEC VAX 11/780-based program, e.g., Hampel et al. (1982). The various and sundry programs deal with pre-processing, i.e., a common command language, and post-processing, i.e., abstract sorting, word processing, etc. Both aspects are being considered in the present program, but with emphasis on the personal computer software. The reasons for this emphasis are:

- i. The rapid increase in micro-computer capability, e.g., the new IBM AT supports 3 mb of main memory and 40 mg of hard disk memory--all in a desk top unit.
- ii. The cost trends. Micro-computer costs have fallen while communications costs have increased. This factor favors distributed computing when compared to a central, i.e., gateway, computer.
- iii. The availability of micro-computer software. There is no question that the number of technology transfer oriented programs has increased many-fold in the last year. This is a strong indication of the market trends.

Many of the tasks associated with conducting technology transfer studies are rather mundane--word processing, producing hard copy of abstracts, etc. While these tasks are fairly straight-forward, they are time consuming. A number of commercial software packages can be used effectively in these tasks.

As noted above, the carrying out of technology transfer studies is labor intensive and software systems can significantly improve productivity. A number of appropriate software packages are shown in Figure 2. These are divided into the general categories of pre-processors, gateways, citation analyzers, report writing and text processors. A short description of each program is given in Table 1. There are over-lapping functions among the software systems as well as the lack of any integration among modules.

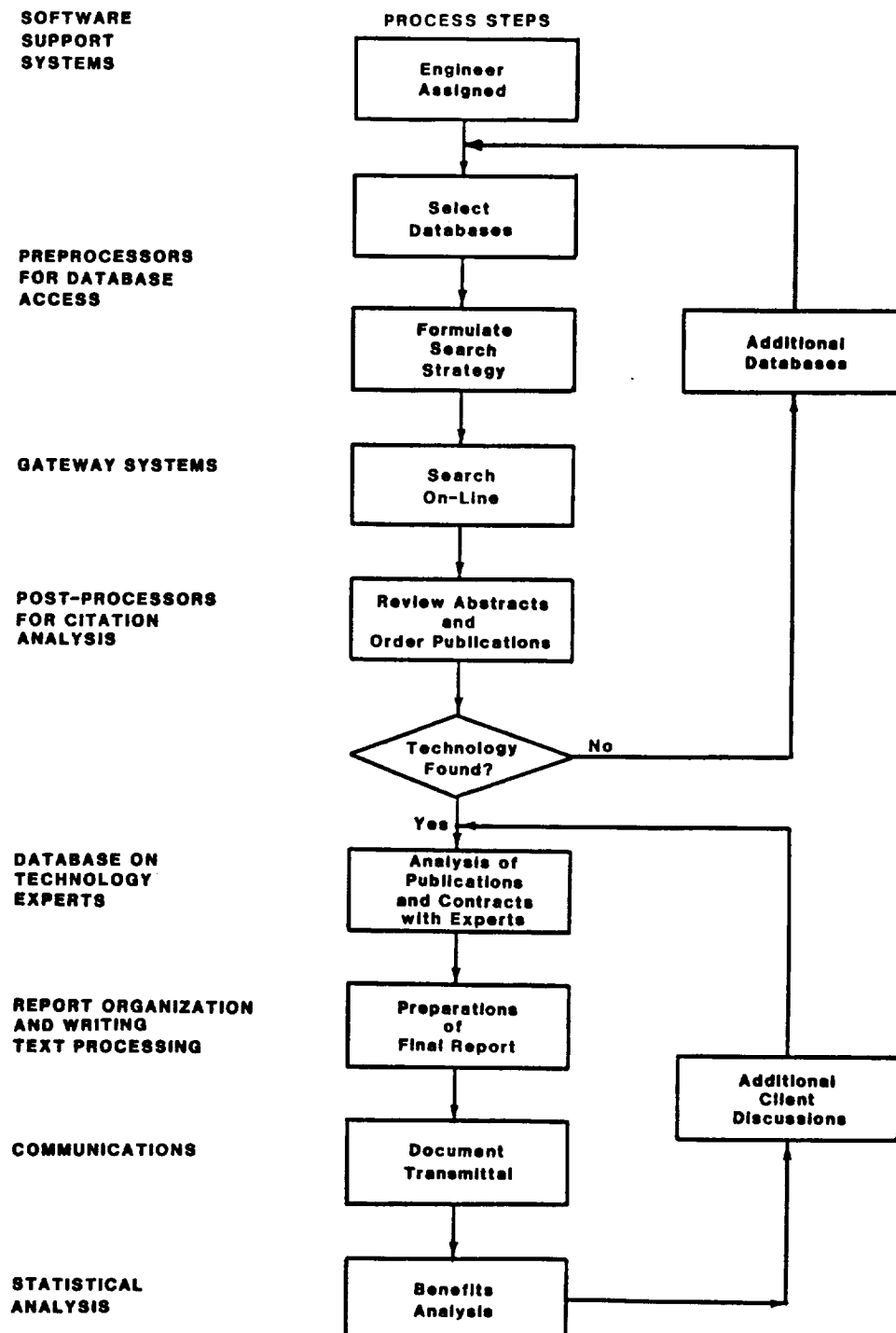


Figure 1. Shown here is a sequence and profile of steps carried out in a typical Industrial Applications Study.

SOFTWARE SYSTEMS		FUNCTIONAL DATA
IN-SEARCH	}	Pre-processors for database access
MIST +		
Business Computer Network	}	Gateway Systems
TOTAL INFORMATION SYSTEM		
NUTSHELL	}	Post-processor for Citation Analysis
dBASE III	}	Database Program
OFFIX	}	Report Organization and Writing
THINKTANK		
MEMORY/SHIFT	}	Integration
TOPVIEW		
WORDSTAR	}	Text processing
SELECT WRITE		
CONNEXUS	}	Communications
STAT-PAK	}	Statistical Analysis

Figure 2. Typical Commercial Software Packages for Use in Technology Transfer

TABLE 1
SYSTEMS RELEVANT TO TECHNOLOGY
TRANSFER TASKS

Program Title	Description	Environment(1)
IN-SEARCH	Front-end database searching program. Catalog of all DIALOG databases and resources. Search revision capabilities and keyword high-lighting within each reference.	IBM PC and compatibles. 192 K; two disk drives; smart or acoustic modem.
MIST +	Microcomputer communications program. Program contains a full programming language with specifications for telecommunicating, a database system and a text editor. The database can be turned into a full-fledged computer teleconferencing system complete with electronic mail, conferences and on-line databanks.	IBM PC and compatibles. 256 K; two disk drives; hard disk recommended; smart or acoustic modem.
Business Computer Network	Logs on automatically to a number of on-line information systems. Program captures text on disk, writes messages off-line, sends them on-line and sends sequences to a printer.	IBM PC and compatibles. 128 K; two disk drives; Hayes compatible modem.

(1) Minimum equipment requirement. The software vendors are listed in Section 6. 128K denotes 128 kilobytes of main memory. IBM PC/XT or compatibles denotes a 16 bit microcomputer having a 10 Mb fixed disk drive. One disk drive denotes a 320-360Kb removable disk drive.

TABLE 1 (continued)
 DESCRIPTIONS OF SOME SOFTWARE
 SYSTEMS RELEVANT TO TECHNOLOGY
 TRANSFER TASKS

Program Title	Description	Environment(1)
NUTSHELL	Citation analysis system which allows review, categorization, etc. Records can be indexed by title, author, keywords, etc.	IBM PC or compatibles. 128 K; one disk drive; smart or acoustic modem.
THINKTANK	Text processing program oriented toward report organization. Uses outlining techniques.	IBM PC/XT or compatibles 256 K; one disk drive.
OFFIX	Personal office system which mimics a file cabinet. Software can search a datafile for up to 10 fields simultaneously and then sort by one of the ten. Can also send information to a screen or printer.	IBM PC/XT or compatibles. 256 K.
MEMORY/SHIFT	This system allows you to run several programs simultaneously and to transfer data between them.	IBM PC/XT or compatibles. 128 K.
TOPVIEW	Provides multitasking and windowing.	IBM PC/XT or compatibles. 512K recommended; one fixed and one removable disk drive.
dBASE III	Database program which constructs and manipulates numeric and character data. Provides database manipulation directly from a keyboard. Provides capability for user-generated menus and application programs.	IBM PC or compatibles. 256K; two disk drives.

TABLE 1 (continued)
 DESCRIPTIONS OF SOME SOFTWARE
 SYSTEMS RELEVANT TO TECHNOLOGY
 TRANSFER TASKS

Program Title	Description	Environment(1)
WORDSTAR	Screen oriented word processing system featuring integrated printing. Displays both initial entry of text and alteration of previously entered text.	IBM PC or compatibles. 128 K; one disk drive.
CONNEXUS	Communications system providing electronic mail, bulletin boards, teleconferencing, etc. Includes password access system.	IBM PC/XT or compatibles. 256 K; two disk drives; smart modem.

1.3 Software Framework for Technology Transfer

Section 1.1 discussed the general steps carried out in technology transfer studies while Section 1.2 reviewed some relevant commercial software systems. This section describes both the framework and elements of a software support system for technology transfer professionals. This system is called a "searcher's workstation" and consists of a series of productivity tools and modules. Commercially available programs, i.e., modules, are being used where possible with unique programs developed only when no suitable systems exist. As discussed in Section 1.2, the software support system is being developed for use on a personal computer--not without precedent, e.g., Bertrand (1980) or Lefkovitz (1982).

Following the steps of Figure 1, the searcher needs a user facility which provides access to the online systems via telephone communications so that he can run searches, use the other features of the online system, and download citations. This need can be met by any of a large number of commercial and public-domain "terminal-emulation programs" which can run on the searcher's local computer.

In addition to this basic capability, it is desirable that the various online database systems be accessible through one common language for specifying searches, downloading citations, and manipulating the resulting files on his local computer. It is possible to obtain this common interface with a program running on the user's local computer. CONIT, Marcus

(1982), is the only program which we have found which meets this need adequately, but CONIT does not seem to be moveable from its present, single (large computer) installation. IN-SEARCH is a commercial offering which works with only some databases and is not user-extendible.

A large number of citations can result from a single search. It is sometimes a large task to classify these citations as to relevancy or sub-category, e.g., Marcus, (1978),. The TIS GATEWAY, Hampel, et al, (1982), offers some capability in this area in the form of tools for manipulating the abstracts and classifying them. In addition, the TIS GATEWAY offers some unique tools for the analysis of citations by time of publication. However, USM has found no commercially available automatic or semi-automatic programs for this purpose.

Report organization and preparation is a major activity. In the area of writing aids, the offerings on the personal computers have out-paced anything before seen. Outlining, organizing, and word-processing tools abound. The searchers should certainly have this capability. There is the necessity of converting the files of citations as down-loaded from the on-line system into a format which can be included into the word-processor. Some word-processors have facilities for doing this; others do not, and it must be programmed. Ultimately, work in this area may lead to some type of automatic presentation of search results.

2. TECHNICAL APPROACH

2.1 Previous Work

2.1.1 User Facility and Common Command Language

A number of faculty members at the University of Southern Mississippi have experimented with a PCCL (Prototype Common Command Language) program. This is a communications program which allows standard database operating commands to be cataloged and called up as macros. The macros are to be supplied by the searcher/user. This program has a menu-based interface. The searcher also designs his menus.

The work with the PCCL program has led to the development of a database user facility program called DBUF. This program is a combination of a structure editor and a communications program, and allows the user to follow and execute expert-supplied scripts for on-line database access.

PCCL and the subsequent DBUF are efforts to provide an interface vehicle for online databases which can be tailored to the expertise level of the user, e.g., Carroll, (1984). This is a more general approach than a "Common Command Language". However, both the PCCL and the DBUF programs can serve as delivery systems for prototypes of such a common command language.

2.1.2 Reviewing and Determining Relevancy of Abstracts

Previous activities have led to the development of SORT-AID which supports online reviewing of down-loaded abstracts. Classification categories can be supplied interactively by the searcher.

Abstract relevancy determination tools are being developed. RANK uses lexical association methods from automatic indexing to rank abstracts. REVIEW is used to facilitate manual reviewing and classification of abstracts. These tools allow relevancy determination to be carried out in a manual, semi-automatic or automatic mode, with more or less reliance on the computer, as the searcher chooses, and as his particular search justifies.

2.1.3 Translation of Citation Formats

The SORT-AID system contains a PRINT module. PRINT supports selection and reformatting of abstracts for printing. Reformatting includes re-justification, re-pagination, and transformation of keywords and content.

2.2 SORT-AID Post-Processor

The SORT-AID system is a set of programs designed to be used by an engineer/researcher in the preparation of industrial applications studies. The SORT-AID system is designed to be useful in that interim stage in report preparation after the database querying and searching have been completed and before the actual writing process begins. Accordingly, SORT-AID provides facilities for classifying, manipulating, and organizing the abstract/citation files resulting from online database searching.

SORT-AID 2.1 is a set of four programs written in FORTRAN 77. SORT-AID 2.1 is a complete re-coding of SORT-AID 2.0. This re-coded version includes all of the features of the older version. In addition, it allows the user to delete citations,

create citations or notes, and reorder the citations within the file. Major improvements in disk storage utilization and execution speed are achieved with this new version.

SORT-AID 2.1 is demonstrated to be portable by implementations on the DEC VAX 11-78, IBM-PC, and UNIVAC, e.g., Vital (1985).

2.2.1 Program NABST

Figure 3 shows the system flow for the SORT-AID system. Input is the individual files resulting from multiple database searches. These files are combined by the NABST program. After the combined file is created by NABST, it can be accessed immediately by the user with the REVIEW program. Optionally, an automatic or semi-automatic relevance ordering can be created for the combined citation file with the RANK program. After the abstract/citations in a combined file have been classified by the user into report-specific categories, they can be printed according to those categories with PRINT. NABST must be run for each abstract/citation file which has been downloaded from a bibliographic database system.

In addition to the abstract/citation file, NABST requires for input a "DELIM.DAT" file. This file is created with the system editor and contains images of the delimiters which may appear in the input abstract/citation files to mark the beginnings of abstract/citations. Furthermore, in addition to the combined abstract/citation file, NABST creates a lexical statistics file which is used by RANK.

2.2.2 Program REVIEW

REVIEW reads the combined abstract/citation file and allows the engineer/searcher to review the contents in its created order or according to a "relevance" order as determined by the RANK program. Optionally the abstracts are presented in the format they were received in (at the request of the users of the system).

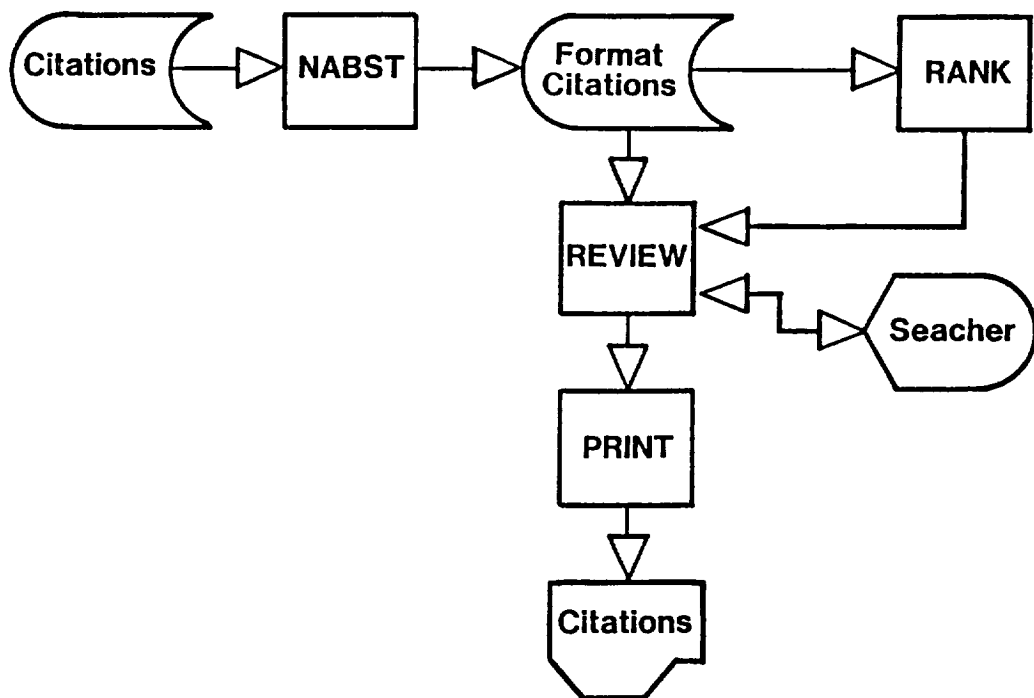


FIGURE 3. System Flow of SORT-AID

REVIEW is command-driven with the following facilities offered:

- i. Proceed to next screen, which will be more of a multi-screen entry, or to the next entry.
- ii. Set abstract category, which allows the operator to assign a report-specific category or categories to be carried in the file with the citation. This category can be any character string without spaces or commas. Multiple categories are entered with commas in between and no spaces.
- iii. Go directly to the next abstract.
- iv. Go to the beginning of the current abstract.
- v. Go directly to an abstract in the file addressed by its relative position in the file.
- vi. Back up one abstract in the file.
- vii. Search on the occurrence of an entered string. This is useful for finding entries containing a particular word, phrase, or author name, for example.
- viii. Set search category; this allows the user to only see entries which have been previously set to a certain category or categories; "*" shows all categories, including un-categorized entries.
- ix. Delete an entry from the file.
- x. Enter notes into the file as a discrete entry; this entry can be categorized and processed like the others.
- xi. Re-order an entry; this may be used to place a particularly interesting entry at the beginning of the file; or an uninteresting one at the end; this is done for the physical and the "relevance" orders.
- xii. Load an internal memory register with the relative position of the current abstract.
- xiii. Return to the abstract pointed to by the internal memory register.
- xiv. Stop processing.

REVIEW is designed to allow the searcher to review and categorize abstracts until he has enough for his report, at which time he can stop processing. It also allows categorization to proceed iteratively via stepwise refinement; that is, the searcher can go through the file once applying gross categories, and then begin at the beginning looking at only a single gross category, applying finer categories.

2.2.3 Program RANK

RANK may be executed after the combined abstract file is completely created by NABST. RANK creates a file which contains a record for each abstract in the combined file. Each record contains the relative position of its associated abstract. These records are ordered in decreasing order by a "relevance" score.

To create the "relevance" score, RANK produces three sets of fifty words each and displays them to the user. The user then specifies the relevance of each word and RANK calculates a "relevance" score.

In addition to the combined citation file, RANK requires a "stop list" file as input. This file contains common words which cannot have report-specific meaning such as "a", "an", "and", and so forth. This file is created with the system editor.

2.2.4 Program PRINT

PRINT is used to include abstracts of particular categories in a report. The format of the abstracts may be changed, i.e., Chen (1985), and the page length and width may be specified. The abstracts are included in the "relevance" order or in physical

order. PRINT is invoked by a command in the REVIEW program. Selection of abstracts for inclusion in the PRINT report follows the conventions set up in REVIEW.

2.3 Abstract Relevance Determination

Abstract relevance determination is a unique requirement of a searcher's workstation. The optimal approach is to develop several tools which the searcher can choose according to his/her tastes and needs. One tool is the REVIEW program in SORT-AID, which allows the examination of each citation and a manual ranking/classification to be done.

Integrated with REVIEW in the SORT-AID system is the RANK program. RANK can order the abstracts according to its own measure of relevance. The searcher can then REVIEW these abstracts in the order, most relevant first, supplied by RANK.

2.4 Theory of Relevance Determination

The RANK program uses a method for semi-automatic relevance determination based on the lexical association methods of automatic indexing, i.e., Salton, (1975). This automatic indexing theory is applied to the abstract/citations which are already the result of a search, rather than to a collection as a whole. This "post-search" collection is comprised of abstract/citations of a homogeneous nature, as it is the product of a single query or related set of queries. This homogeneous nature of the collection accentuates the characteristics of the automatic indexing methods. The RANK program exploits this accentuation.

The method of relevance ranking employed may be considered to be similar to the keyword querying. However, the keywords are not selected from a controlled vocabulary of the database system, but from lists generated with automatic indexing algorithms applied internally to the "post-search" collection. The user is allowed to select his "query" keywords from these lists. The abstracts in the collection are ranked for relevance by the frequency of occurrence of these query keywords.

The lexical association metrics used for "keyword" generation are based on automatic indexing theory, e.g., Salton (1975a), (1975b), (1968), Salton and McGill (1983) and von Rijsbergen (1980). In particular, the frequency and/or number of occurrences of terms within abstracts are determined. This data takes the form shown in Figure 4 with $n_{j,k}$ denoting the number of occurrences of the k -th term in the j -th abstract. The frequency of the term k in the collection can be determined as

$$F_k = \sum_{j=1}^J n_{j,k} \quad (1)$$

Note that frequency and number of occurrences are used more or less interchangeably in automatic indexing.

F_k can be biased by a few very large values of $n_{j,k}$. A document occurrence and/or frequency can be defined as

$$b_{j,k} = \begin{cases} 1 & n_{j,k} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

and

$$B_k = \sum_{j=1}^J b_{j,k} \quad (3)$$

TERM	ABSTRACT				
k	j:	1	2	3	---
1		n _{1,1}	n _{2,1}	n _{3,1}	---
2		n _{1,2}	n _{2,2}	n _{3,2}	---
3		n _{1,3}	n _{2,3}	n _{3,3}	---
-		-	-	-	---
-		-	-	-	---
-		-	-	-	---
K		n _{1,K}	n _{2,K}	n _{3,K}	---
					n _{J,K}

Figure 4. Occurrence Matrix, $n_{j,k}$

An additional factor describing term behavior is the signal-noise ratio, e.g., Shannon (1948). The signal and noise are computed for the complete collection of abstracts and

$$N_k = 1.4427 \left\{ \ln F_k - \frac{1}{F_k} \sum_{j=1}^J n_{j,k} \ln(n_{j,k}) \right\} \quad (4)$$

where

$$\begin{array}{l} n_{j,k} \ln(n_{j,k}) \xrightarrow{\quad} 0 \\ n_{j,k} \xrightarrow{\quad} 0 \end{array}$$

by L' Hospital's rule. The signal is defined as

$$S_k = 1.4427 \ln F_k - N_k$$

Note that the factor 1.4427 results from the conversion of $\log_2 X$ to $\log_e X$, i.e., $\ln X$.

Some insight in the behavior of S_k and N_k can be gained by considering the case where $n_{j,k} = 1$ for all j and a specified k . It follows that $F_k = J$ and $N_k = 1.4427 \ln(J)$ with $S_k = 1.4427 \ln J - 1.4427 \ln J = 0$. In this particular case, the noise level is constant with no signal. Consider the case where all $n_{j,k} = 0$ save one having the value F_k . It then follows that $N_k = 0$ and $S_k = 1.4427 \ln F_k$. These represent the limiting cases.

The three parameters which describe the term distributions are F_k , B_k and S_k . These can be used to evaluate the importance of the terms in the following manner. The weighting factors $w_k^{(f)}$, $w_k^{(s)}$ and $w_k^{(b)}$ can be determined as

$$w_k^{(f)} = F_k D_k \quad (6)$$

$$w_k^{(b)} = B_k D_k \quad (7)$$

$$w_k^{(s)} = F_k D_k S_k \quad (8)$$

where D_k is defined as

$$D_k = 1.4427 [\ln(J) - \ln(B_k) + 1] \quad (9)$$

with the various terms ranked according to the w_k values. The first 50 words are displayed with a "relevance" value, λ_k , assigned to each word. The relevance value can be considered with the term characteristics to rank the abstracts according to relevance.

The ranking equations are generated by summing over the term index and

$$R_j^{(f)} = \frac{1}{K} \sum_{k=1}^K n_{j,k} D_k \lambda_k^{(f)} \quad (10)$$

$$R_j^{(s)} = \frac{1}{K} \sum_{k=1}^K n_{j,k} D_k S_k \lambda_k^{(s)} \quad (11)$$

$$R_j^{(b)} = \frac{1}{K} \sum_{k=1}^K b_{j,k} D_k \lambda_k^{(b)} \quad (12)$$

where R_j denotes the relative ranking of the j -th abstract. The superscript f , s or b simply denotes the method.

2.5 Evaluation of Relevance Determination

The evaluation of relevance is of a subjective nature and will vary with the evaluator. The usefulness of the approach will be determined by the users. This will necessitate field testing and modification which is described in Section 2.7. In an attempt to conduct a preliminary evaluation, a set of 164 abstracts has been ranked using all three methods. The abstracts were the result of an industrial applications study which was conducted at the Aerospace Research Applications Center. The abstracts were categorized with 33 out of 164 being judged relevant, i.e., 20% which is typical.

The results of the preliminary evaluation are shown in Figure 5, 6, 7, 8, and 9. Figures 5, 6, and 7 list the 50 words and related relevance, i.e., W_k and λ_k , for the three methods. Figures 8 and 9 depict the fraction of relevant abstracts by 1/8 and 1/4 groups. AR denotes as received, FK ranking using $R_j^{(f)}$, SK ranking using $R_j^{(s)}$ and BK ranking using $R_j^{(b)}$. Note that 33 of the 164 abstracts were judged relevant. The ideal distributions are denoted as ID and are also shown. Note that SK and FK significantly improve the distribution, i.e., concentrate the relevant abstracts in the first groups, but are far from ideal. Additional activities in this area are envisioned.

2.6 Proposed Continuation of Work

The major activity during the current contract year, i.e., July 1, 1985 to June 30, 1986, has consisted of converting SORT-AID from a DEC VAX 11/780 operating mode to a system which

functions on an IBM PC/XT or equivalent. This has proved to be more difficult than originally envisioned and has required a complete recoding of the entire program. As a result, field testing has been delayed with installations at the four test sites now scheduled for the mid-March to mid-April time period. User experience will be generated during May to July with program modifications made in the August - October time frame.

$\lambda_k(f)$	Term	$w_k(f)$	$\lambda_k(f)$	Term	$w_k(f)$
10	PUMP	1084	0	CENTRIFUGAL	186
10	PRESSURE	647	0	WATER	185
10	HYDRAULIC	609	10	LIQUID	173
0	SEAL	569	0	MEASUREMENT	171
10	SPEED	541	0	STAGE	169
0	FLUID	490	0	OILHYDRAULIC	167
0	CIRCUIT	450	0	BHRA	166
10	NOISE	449	0	IMPELLER	164
10	GEAR	432	20	TRANSMISSION	160
10	POWER	430	0	DEVELOPMENT	159
0	FLOW	390	10	EFFICIENCY	157
10	VALVE	319	0	ENGINE	157
10	VANE	315	0	DEVICE	156
10	RANGE	260	0	TRANSPORT	156
0	TURBINE	259	0	COMPONENT	155
10	DISPLACEMENT	229	0	PROPULSION	155
10	ROTARY	207	0	CHARACTERISTIC	153
10	OIL	206	0	AIR	153
10	DRIVE	205	10	VIBRATION	149
10	VARIABLE	203	0	STANDARD	147
0	FACE	200	10	VORTEX	137
10	PISTON	195	0	CAVITATION	137
0	COUPL	189	0	REQUIREMENT	136
0	SOLID	187	0	MADE	134
10	PERFORMANCE	186	0	RING	133

Figure 5. Keywords for Abstract Relevance, $w_k(f)$

$\lambda_k(s)$	Term	$w_k(s)$	$\lambda_k(s)$	Term	$w_k(s)$
10	PUMP	2648	0	DEVICE	260
0	SEAL	1644	10	RING	243
10	PRESSURE	1257	0	CENTRIFUGAL	241
10	NOISE	1110	10	VARIABLE	233
0	CIRCUIT	1051	10	VIBRATION	227
10	SPEED	1013	0	AIR	217
10	HYDRAULIC	1002	0	WATER	210
10	GEAR	808	10	PISTON	210
0	FLUID	689	0	IMPELLER	208
10	POWER	600	10	LIQUID	207
0	FLOW	593	0	MEASUREMENT	199
10	VANE	545	0	THRUST	195
10	VALVE	509	10	ROTARY	193
0	TURBINE	507	20	FLUIDIC	190
10	VORTEX	380	0	CAM	187
0	PROPULSION	380	0	DRIVE	187
0	COUPL	358	20	HYDROSTATIC	183
0	FACE	354	0	BHRA	182
0	SOLID	330	10	TORQUE	179
0	ENGINE	321	0	DEVELOPMENT	176
0	WATERJET	306	0	LIFT	176
0	STAGE	291	0	BEAR	174
10	OIL	278	0	PACK	161
0	TRANSPORT	263	0	SAND	160
10	DISPLACEMENT	261	0	HITACHI	154

Figure 6. Keywords for Abstract Relevance, $w_k(s)$

$\lambda_k(b)$	Term	$w_k(b)$	$\lambda_k(b)$	Term	$w_k(b)$
10	PUMP	236	0	ENGINEER	101
0	FLUID	231	0	PROC	101
10	HYDRAULIC	227	10	OIL	101
10	PRESSURE	223	0	WATER	101
10	POWER	203	0	MADE	101
10	SPEED	202	10	NOISE	98
10	RANGE	199	0	WIDE	98
0	FLOW	185	10	LIQUID	95
0	OILHYDRAULIC	161	20	TRANSMISSION	95
10	GEAR	143	0	CONF	92
10	PERFORMANCE	132	0	DISCUSS	92
10	VALVE	132	0	AXIAL	89
10	DRIVE	128	0	STANDARD	89
10	ROTARY	126	0	ONLY	89
0	CIRCUIT	123	0	CENTRIFUGAL	86
10	DISPLACEMENT	123	0	DUE	86
0	COMPONENT	121	0	REQUIREMENT	86
10	VARIABLE	119	0	DEVELOPMENT	86
0	CHARACTERISTIC	109	0	NECESSARY	86
10	VANE	109	0	PRODUCT	86
10	EFFICIENCY	109	0	CLAIM	86
0	MAXIMUM	106	0	BHRA	83
0	GERMAN	103	0	PROVIDE	83
0	SEAL	103	0	MEASUREMENT	83
10	PISTON	103	0	TURBINE	83

Figure 7. Keywords for Abstract Relevance, $w_k(b)$

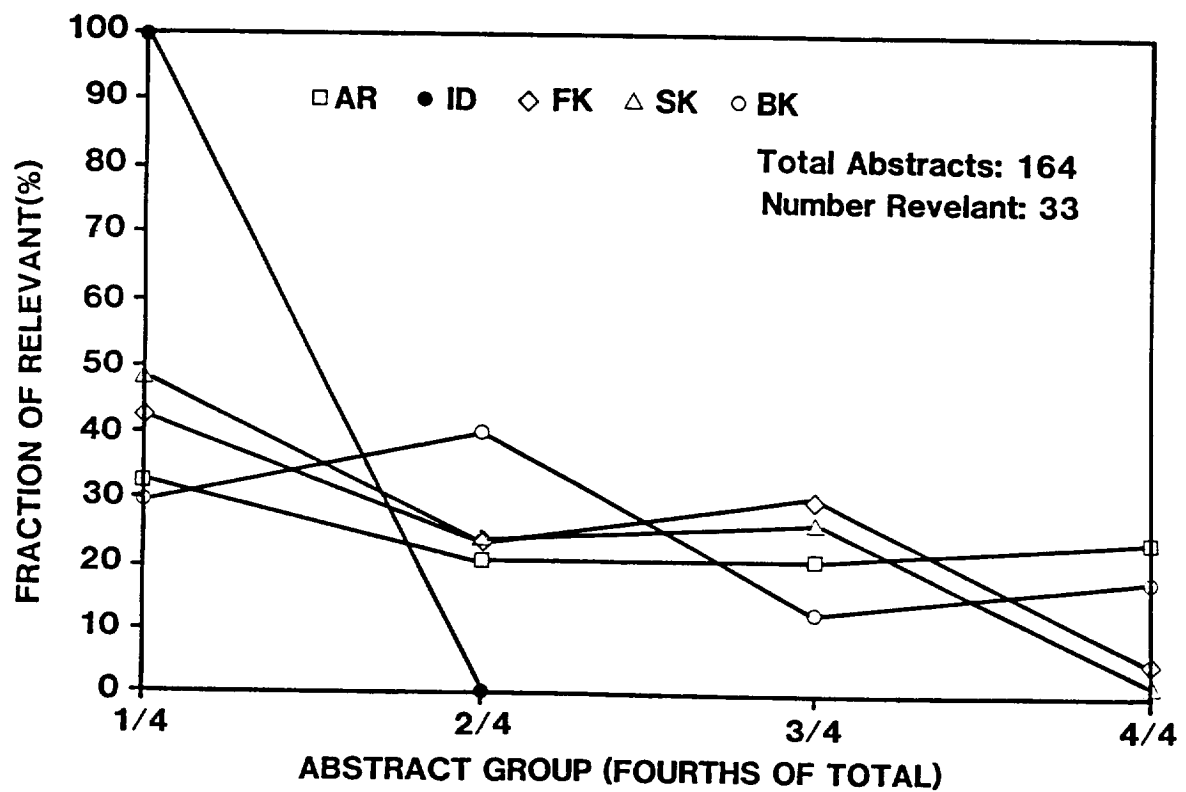


Figure 8. Abstract Distribution by Groups

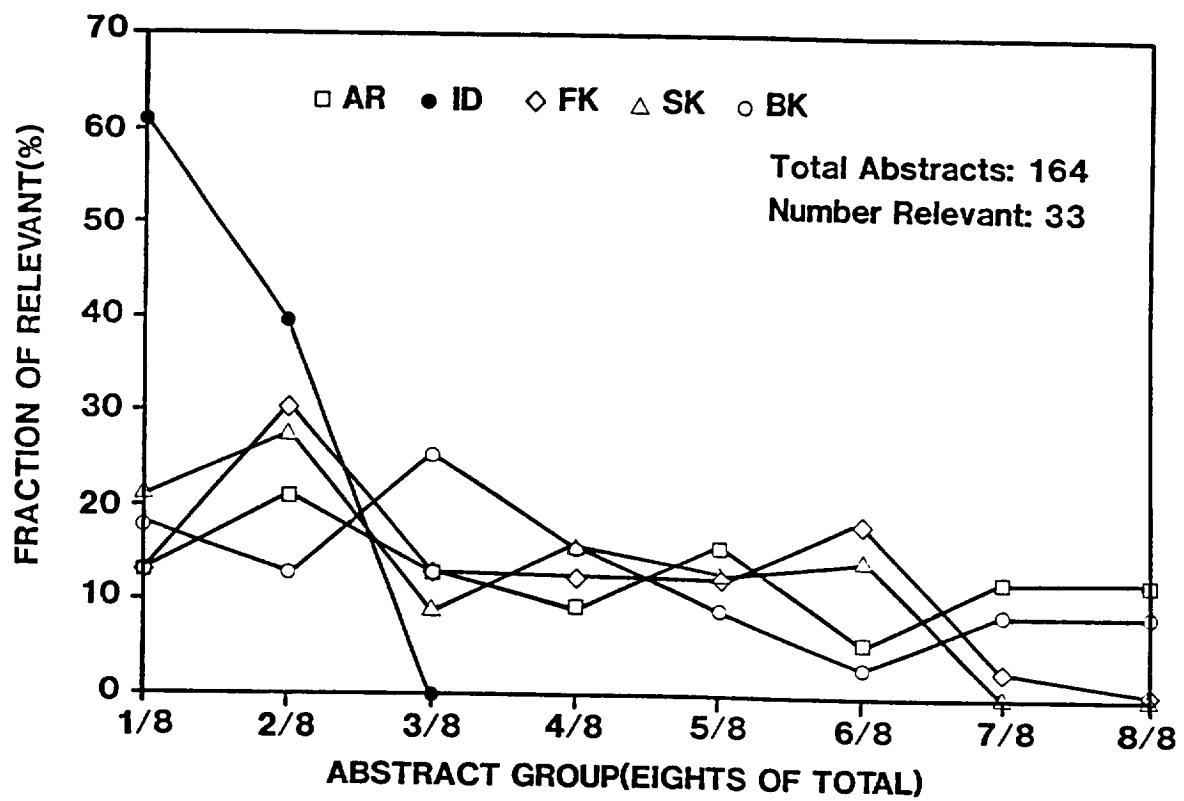


Figure 9. Abstract Distribution by Groups

The program modifications--as currently envisioned--deal with the two major elements of the project--RANK and REVIEW. Recall that REVIEW is the "manual" abstract processor while RANK provides "semi-automatic" relevancy determination. Operating experience with REVIEW and to a lesser extent with RANK has been accrued at the Aerospace Research Applications Center (ARAC). Note that ARAC operated the software on a DEC VAX 11/780. ARAC's experiences to date are documented by Goehring (1986) and are positive. ARAC made a number of additions to the program--command files, etc.--and these will be incorporated during the program modification phase. Other suggestions, e.g., sorting by reverse chronological order, automatic generation of bibliographies, elimination of duplication, etc. will be considered for implementation. A number of suggested changes are likely to arise during testing and these will be evaluated and incorporated if deemed worthwhile and feasible.

Operating experience with RANK is very limited. The micro-computer version will initially provide three separate lexical association techniques which are user selectable. Testing to date indicates that the frequency and signal to noise approaches are effective. Testing at USM will continue on these methods during the current contract year. Additional approaches will be evaluated during the continuation period, i.e., July 1, 1986 to June 30, 1987.

In particular, a fourth lexical association method based on the variance will be evaluated. The variance of term occurrence, V_k^2 , can be defined as

$$(V_k)^2 = \frac{1}{J-1} \sum_{j=1}^J (n_{j,k} - \bar{n}_k)^2 \quad (13)$$

where

$$\bar{n}_k = \frac{1}{J} \sum_{j=1}^J n_{j,k} = \frac{F_k}{J} \quad (14)$$

The variance can be written as

$$(V_k)^2 = \frac{1}{J-1} \sum_{j=1}^J (n_{j,k})^2 - \frac{F_k^2}{J(J-1)} \quad (15)$$

Weighting and ranking factors can then be defined as

$$W_k^{(v)} = F_k D_k V_k \quad (16)$$

and

$$R_j^{(v)} = \frac{1}{K} \sum_{k=1}^K n_{j,k} D_k V_k \lambda_k^{(v)} \quad (17)$$

At the present time, the methods of equations (6), (7), and (8) are being applied independent of one another. Previous approaches, e.g., Huffman, et. al. (1986), have combined techniques. For example, the frequency and signal-to-noise approaches were combined to produce an abstract ranking. This approach is difficult to implement on a micro-computer but will be considered during the coming year.

The results of Figures 5, 6, and 7 rely heavily on the relevance factor λ_k for the ultimate abstract ranking. An alternate approach is to compare all abstracts with a relevant

baseline document. The baseline document could be one member of the abstract set or a problem statement. The abstracts would then be ranked on the comparison metric. Consider the parameter

$$\beta_{j,k} = n_{j,k} D_k S_k \quad (18)$$

where $n_{j,k}$, D_k and S_k have the definitions of Section 2.4. A similarity measure, $s_{i,j}$ can be defined as

$$s_{i,j} = \sum_{k=1}^K \beta_{i,k} \beta_{j,k} \quad (19)$$

where the subscripts i and j denote the i -th and j -th document respectively. If I denotes a reference document, then $s_{I,j}$ denotes a similarity measure between I and j . The documents can then be ranked as

$$R_j^{(\beta)} = s_{I,j}$$

with the abstracts with the largest R_j values being the most relevant.

It is important to recognize that the abstract relevance determination is based on largely observational data. The approach combines linguistics and statistics along with some general characteristics of the scientific use of the English language. Unlike the analysis of physical systems, there are no physical laws, i.e., conservation of mass, energy, etc., which can be applied. As a result, experimentation is the rule of the day. Much trial and error is required. This does not imply a potential lack of success but rather necessitates a thorough evaluation with liberal input from the user community. Results

to date are encouraging and the methods possess tremendous labor savings potential.

2.7 Scope of Work

2.7.1 Work Statement

The NASA Technology Utilization office in conjunction with the University of Southern Mississippi has supported the development of a number of software tools which aid in the practice of technology transfer. The technology transfer process includes problem definition, bibliographic database searching, analysis of abstracts and/or full text documents, consultation with relevant experts and preparation of a final report. The software tools support all of these activities to a greater or lesser extent with primary emphasis on the analysis of abstracts. The continuation of activities is principally focused on completion of the abstract analysis systems known as SORT-AID. SORT-AID determines the relevance of abstracts using a series of lexical association techniques.

The basic items of work to be accomplished in the continuation program are:

- i. Software installation and support. The SORT-AID software system will be installed in two additional sites, i.e., NASA industrial application centers (IAC) or state technology assistance centers (STAC). USM will provide software support for these operating organizations during the test phase.
- ii. Continuing optimization of micro-computer based SORT-AID. SORT-AID has been completely recoded in the DEC VAX 11/780 translation process. The program is currently operational and fully-functional. It is believed, however, that computing times can be reduced by code optimization.

- iii. Continuing evaluation of abstract relevancy approaches. As noted in Sections 2.4 and 2.5, the evaluation and/or modification of the lexical association methods is absolutely imperative. New techniques will be developed and evaluated at USM prior to field testing.
- iv. Continued software testing. The prototype software system will be installed at three NASA sites during the current contract year. The testing will continue into the proposed contract year with USM providing complete software support. The test facilities will be expected to monitor system performance, log and/or record problems and generally offer suggestions with regard to the system performance.
- v. Software modifications. Following reports and/or comments by the three test sites, revisions to the software system will be made if required.
- vi. Modifications to the software documentation. All software furnished to the NASA IAC's and STAC's will be user-friendly and documented to a standard compatible to commercially-available software. The software will be modified to be consistent with any functional changes. Furthermore, user input will be reviewed with regard to improving readability, understandability, etc.
- vii. Liaison with other NASA Information/Technology Transfer Organizations. USM personnel will continue liaison activities with other NASA funded programs on a time-available basis.
- viii. Reporting. Three quarterly reports and a final report will be provided.

2.7.2 Schedule

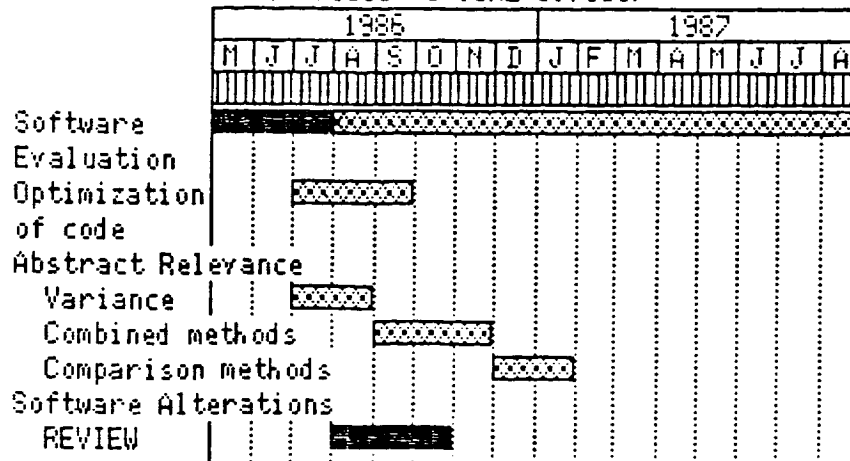
The schedule for the aforementioned tasks is shown in Figure 10. The schedule is realistic and is consistent with the funding available.

2.7.3 Budget

The program budget is shown in Table 2. The total program cost is \$84,461 with USM providing \$11,474 in cost sharing. The net cost to NASA is \$72,987.

SOFTWARE DEVELOPMENT FOR THE NASA TECHNOLOGY UTILIZATION PROGRAM

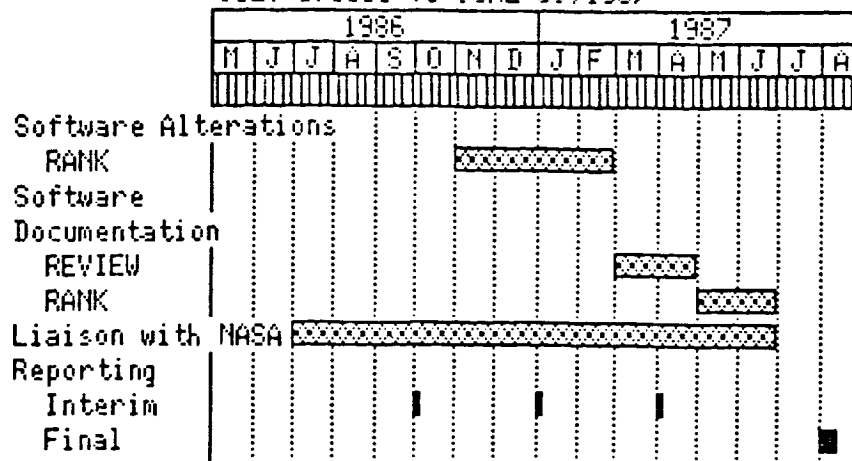
JULY 1, 1986 TO JUNE 30, 1987



Solid bars denote continuous activity while cross-hatched bars denote intermittent activity.

SOFTWARE DEVELOPMENT FOR THE NASA TECHNOLOGY UTILIZATION PROGRAM

JULY 1, 1986 TO JUNE 30, 1987



Solid bars denote continuous activity while cross-hatched bars denote intermittent activity.

Figure 10. Program schedule

Table 2
Program Budget

<u>Personnel</u>	NASA	USM	Total
G. David Huffman, 3.6mm	15,000	5,020	20,020
G. Oehms, 2mm	3,000	2,000	5,000
D. Vital, 5mm	10,000	0	10,000
Research Assistant, 6 mm	6,000	0	6,000
Part-time secretarial, 3.0 mm	3,000	0	3,000
Subtotal	37,000	7,020	44,020
 <u>Fringe Benefits</u>	 5,937	 1,295	 7,232
 <u>Miscellaneous</u>			
Software	400	0	400
Travel	8,000	0	8,000
Subtotal	8,400	0	8,400
 <u>Equipment</u>	 5,000	 0	 5,000
 <u>Indirect Costs</u> (45% of salaries)	 16,650	 3,159	 19,809
 <u>Total</u>	 72,987	 11,474	 84,461

2.7.4 Summary

Technology transfer activities are carried out by NASA via a network of Industrial Applications Centers and State Technology Assistance Centers. These organizations employ a number of modes of technology transfer which principally vary in degree of specificity. These modes are largely the same regardless of the technology under consideration. The processes are labor intensive and are now being automated using micro-computer-based techniques and software support systems. The systems are currently being installed with testing, optimization and modification proposed for the coming contract year. Full implementation of the software support systems can reduce technology transfer study costs by up to 20-25%.

3. PERSONNEL

3.1 Overview

The program will be conducted by Dr. Huffman with the support of Mr. Oehms. Administrative support will be provided by Ms. D. Theisen. Dr. Huffman and Mr. Oehms have over 40 years of combined experience in software development. Their experience has been in both universities and industry and spans a broad spectrum of activities, i.e., real-time software for process control to management information systems. Resumes of all individuals are listed in the following sections.

3.2 Resume for G. David Huffman

PERSONAL DATA

Office: College of Science and Technology
The University of Southern Mississippi
Southern Station, Box 5165
Hattiesburg, MS 39406
Telephone (601) 266-4883

Residence: 902 Sioux Lane
Hattiesburg, MS 39401
Telephone (601) 264-6137

EDUCATION

1970 Postdoctoral research, Imperial College of
Science and Technology, London, England

1968 Ph.D., Mechanical Engineering, The Ohio State
University, Columbus, OH

1966 M. Sc., Mechanical Engineering, The Ohio
State University, Columbus, OH

1962 B. Engr. Sc. (cum laude), Engineering
Science, Marshall University, Huntington, WV

PROFESSIONAL EXPERIENCE

August 1984 - Present Dean of the College of Science and Technology
and Professor of Computer Science and
Engineering Technology, The University of
Southern Mississippi, Hattiesburg, MS.

December 1983 - Professor and Chairman, Department of
July 1984 Computer and Information Science, Purdue
University School of Science, Indianapolis,
IN; Chief Scientist, Indianapolis Center for
Advanced Research; Adjunct Professor of
Engineering, Purdue University School of
Engineering and Technology, Indianapolis,
IN.

August 1974 - Director, Energy Engineering and Research
August 1984 Division, Indianapolis Center for Advanced
Research and Professor of Computer Science,
Purdue University School of Science, Indiana
University-Purdue University at Indianapolis,
IN.

July 1978 - June 1979	Distinguished Visiting Scientist, Department of Aeronautics, U.S. Air Force Academy, CO.
June 1974 - August 1974	Chief, Mechanics Research, Detroit Diesel Allison Division, General Motors Corporation, Indianapolis, IN.
Sept. 1970 - Sept. 1972	Principal Scientist, Experimental Fluid Mechanics Research, Detroit Diesel Allison Division, General Motors Corporation, Indianapolis, IN.
Sept. 1968 - Sept. 1970	Section Chief Research Department, Detroit Diesel Allison Division, General Motors Corporation, Indianapolis, IN.
March 1962 - August 1965	Research Engineer, U.S. Air Force Aero Pro- pulsion Laboratory, Wright-Patterson AFB, OH.

3.3 Resume for Glenn Edward Oehms

PERSONAL DATA

Office: College of Science and Technology
The University of Southern Mississippi
Southern Station, Box 5165
Hattiesburg, MS 39406
Telephone (601) 266-5144

Residence: Post Office Box 524
Hattiesburg, MS 39401

EDUCATION

Sept. 1983 - Received M.S. in Computer Science.
May 1986 University of Southern Mississippi (USM)

Jan. 1982 - Completed 21 graduate hours and 9 under-
graduate hours toward M.S. degree in
Computer Science

August 1981 - Received M.A. from USM--English/Linguistics.

June 1981 - Certificate from Institute in Technical
Communication sponsored by SCETC.

August 1978 - Received B.A. from USM--English.

Sept. 1969 - Defense Language Communication sponsored by
March 1970 SCETC.

August 1967 - Defense Language Institute, Washington, D.C.
August 1968 Hindi-Urdu Language.

EMPLOYMENT HISTORY

Nov. 1985 - Director of Computing Facilities, College of
Current Science and Technology, USM

August 1985 - Graduate Teaching Associate, Dept. of
October 1985 Computer Science, USM

June 1984 - Instructor, Dept. of Computer Science, USM
August 1985

Sept. 1983 - Research Assistant, Dept. of Computer
May 1984 Science, USM

Sept. 1982 - Lab Assistant, CSS3700L, Dept. of Computer
August 1983 Science, USM

Sept. 1979 - USM Faculty; Composition Instructor.
May 1980

Sept. 1978 - USM Teaching Assistant--Project Access
May 1979

Sept. 1977 - Denlinger's Publishing Company, Centreville,
March 1978 VA: Editor

May 1977 - Pas-Point Ambulance, Pascagoula, MS:
August 1977 Emergency Medical Technician/Driver.

August 1974 - Mobile Medic Ambulance Service, Gulfport,
July 1976 MS: Training Officer/Shift Supervisor/
Paramedic/Driver.

3.4 Resume for Dennis A. Vital

PERSONAL DATA

Office: College of Science and Technology
University of Southern Mississippi
Southern Station Box 5165
Hattiesburg, MS 39406

Residence: 208 1/2 N. 22nd Ave.
Hattiesburg, MS 39401

EDUCATION

Present Working to complete M.S. in Computer
Science at USM. Course work finished,
thesis to be completed.

May 1983 Received B.S. in Computer Science from
the University of Southern Mississippi

WORK EXPERIENCE

June 1986 - Systems Analyst, College of Science and
Present Technology Computing Facilities, USM

Nov. 1984 - Research Assistant, College of Science
May 1986 and Technology, USM

Aug. 1979 - USM Food Services, final position: Student
Oct. 1984 Assistant Banquet Manager

3.5 Resume for Deborah A. Theisen

PERSONAL DATA

Office: College of Science and Technology
University of Southern Mississippi
Southern Station Box 5165
Hattiesburg, MS 39406

Residence: 103 Ross Blvd., Apartment A-2
Hattiesburg, MS 39401

EDUCATION

Present Pursuing a degree in Accounting, University
of Southern Mississippi, Hattiesburg,
Mississippi

1975 Diploma - Pascagoula High School, Pascagoula,
Mississippi

WORK EXPERIENCE

August 1985 - Office Manager of the Dean's Office, College
Present of Science and Technology, USM

August 1983 - Secretary, Department of Engineering Tech-
July 1985 nology, University of Southern Mississippi,
Hattiesburg, Mississippi

January 1981 - Teacher's Aide - Special Education, Hatties-
May 1982 burg Public School System, Hattiesburg,
Mississippi

July 1979 - Electrical Apprentice, Ingalls Shipbuilding,
August 1980 Litton Industries, Pascagoula, Mississippi

October 1976 - Secretary, Ingalls Shipbuilding, Litton In-
June 1979 dustries, Pascagoula, Mississippi

4. OPERATING ORGANIZATION

4.1 The University of Southern Mississippi

The University of Southern Mississippi was established by an act of the Legislature approved on March 30, 1910, by Governor Edmund E. Noel. Its first name was the Mississippi Normal College, and its original purpose was to train teachers for the rural schools of Mississippi.

The act of March 30, 1910, did not provide any state money for the building of Mississippi Normal College, but did provide that localities in the state might bid for its location by offering land for a site and money for constructing buildings. On September 16, 1910, the Board of Trustees accepted the bid of Hattiesburg and Forrest County to supply \$250,000 and a free site. That site was west of the city in cut-over timberland with great pine stumps everywhere. Contracts were left to clear the land and to build buildings.

The five permanent buildings (College Hall, Forrest County Hall, Hattiesburg Hall, the Industrial Cottage, now the Honor House, and the President's Home, now the Alumni House, a temporary wooden Dining Hall) and other necessary improvements were barely finished when the Mississippi Normal College opened on the rainy morning of September 18, 1912, with a president, a faculty of eighteen, and a student body of 200.

On October 17, 1911, Joseph Anderson Cook, Superintendent of Schools, Columbus, Mississippi, was elected president. The University of Southern Mississippi has had only six presidents

during the more than seventy years since its founding. The Board of Trustees elected Claude Bennett president effective October 10, 1928. On April 23, 1933, the Board of Trustees elected Dr. Jennings Burgon George as the third president, effective July 1, 1933. On June 13, 1945, the Board of Trustees elected Dr. Robert Cecil Cook as the fourth president and he officially assumed office on July 1, 1945. On October 21, 1954, President Cook submitted his resignation. He served until December 31, 1954 and Dr. Richard Aubrey McLemore became acting president on January 1, 1955. The Board of Trustees, on May 19, 1955, elected Dr. William David McCain as the fifth president. He officially assumed office on August 1, 1955, and retired as of June 30, 1975. Dr. Aubrey Keith Lucas became the sixth president of the University on July 1, 1975.

As has been stated, the University of Southern Mississippi was founded on March 30, 1910, as the Mississippi Normal College. On March 7, 1924, the Legislature changed the name to State Teachers College. On February 8, 1940, the Legislature changed the name to Mississippi Southern College, and on February 27, 1962, the Legislature changed the name to the University of Southern Mississippi.

The Mississippi Normal College did not grant degrees in its early years, but awarded certificates for the completion of certain specified courses of study. On April 8, 1922, the Legislature authorized the awarding of the Bachelor of Science degree. The Bachelor of Music degree was authorized by the Board

of Trustees on June 19, 1934. The first Bachelor of Arts degree was awarded on August 20, 1940. On May 26, 1947, the Board of Trustees of State Institutions of Higher Learning authorized the initiation of graduate work and the awarding of the Master of Arts. Doctoral programs were first authorized by the Board of Trustees on May 20, 1959. Numerous other degree programs have been authorized during the past twenty-six years.

The administrative and academic organization of the University of Southern Mississippi is basic and relatively simple. This reorganization took place in 1982. The Executive Vice President coordinates the offices of the Vice President for Academic Affairs and the Vice President for Research and Extended Services. The Vice President for Business and Finance is in charge of all financial operations of the University. The Vice President for Student Affairs is responsible for the welfare of the student body, and the Director of Intercollegiate Athletics is in charge of the athletic programs.

The academic area of the University of Southern Mississippi is organized into the College of Business Administration, the College of Education and Psychology, the College of Fine Arts, the College of Liberal Arts, the College of Science and Technology, the Honors College, the School of Health, Physical Education and Recreation, the School of Home Economics, the School of Library Service, the School of Nursing, the Graduate School of Social Work, the Graduate School, and the Division of Continuing Education.

The number of graduates is some measure of the growth of a university. The University of Southern Mississippi has awarded 56,606 degrees since the first one was conferred in 1922.

Research has been strengthened at the University of Southern Mississippi in several ways; the number of doctoral degrees has been increased; the University has established its own scholarly journals The Southern Quarterly and The Education and Psychological Research Journal; the Computing Center has been equipped to expedite contracts and faculty research; the Office of Research and Sponsored Programs has been established by the University to promote research and to assist faculty members in their research activities; and the University of Southern Mississippi is a participating member of the University Press of Mississippi established on May 1, 1970.

The University of Southern Mississippi was established by the State of Mississippi, is owned by the State of Mississippi, and is operated and financed by the State of Mississippi. Its first Board of Trustees was established by the legislative act of March 30, 1910, and that board governed only this one institution. On February 2, 1932, the Legislature established the Board of Trustees of State Institutions of Higher Learning and placed under its jurisdiction the five colleges and one university owned and operated by the state. On November 3, 1942, the people of the State voted to make the Board of Trustees of State Institutes of Higher Learning a constitutional board for all colleges and universities of the State. The University of

Southern Mississippi is now operated under the jurisdiction of that constitutional board.

4.2 The College of Science and Technology

The College of Science and Technology provides the student training in all of the classical fields of science, several contemporary multidisciplinary areas, and engineering technology degree programs. The College of Science and Technology is organized into eleven departments and five institutes. The departments include Biological Sciences, Chemistry, Computer Science and Statistics, Construction and Architectural Engineering Technology, Geology, Engineering Technology, Mathematics, Medical Technology, Physics and Astronomy, Polymer Science, and Science Education. The institutes are the Institute of Environmental Science, the Institute of Genetics, the Institute of Microbiology and Related Sciences, the Institute of Surface Coatings, and the Mississippi Polymer Institute. In addition to the degree programs that are synonymous with the above listed department names, degrees are offered in Industrial Engineering Technology, Electronics Engineering Technology, Mechanical Engineering Technology, and Computer Engineering Technology.

Pre-Professional Curricula are offered by the College of Science and Technology in the following health related areas: medicine, dentistry, veterinary medicine, pharmacy, physical therapy, optometry, dental hygiene, and medical records administration. The College also provides a pre-engineering and a pre-architecture curriculum.

The College shares with the College of Education and Psychology joint responsibility for the Department of Science Education, embracing a sequence of courses titled Fundamentals of Science, and offering a major in science education for teachers.

4.3 The Department of Computer Science and Statistics

The Department of Computer Science and Statistics was established in 1965 to meet the great demand for trained personnel in the computing industry. Since 1965, over 1,200 students have obtained degrees in the program. The department currently offers B.S. and M.S. degrees and is contemplating initiating a Ph.D. program.

Majors in the department currently consists of computer science, business data processing, applied computer science, and statistics. Courses offered by the department include programming languages, compilers, operating systems, computer design, microprocessor interfaces, communication networks, process control systems, analog computing, switching circuits, database design, linear and mathematical programming, simulation, statistics, and other miscellaneous topics.

4.4 College Computer Facilities

The primary systems for faculty and student support within the college are a DEC VAX-11/780 and a Harris H-800 computer system. This DEC VAX-11/780 is currently configured with 8 mb of main memory, 512 mb of disk storage, two 8-track tape drives, and a number of miscellaneous peripherals. The system currently supports 80 terminals and can be externally accessed through a

number of "dial-up" lines. The Harris H-800 has 3 mb of main memory, 600 mb of disk storage, one 8-track tape drive and supports 20 terminals and 64 "dial-up" lines. Additional computer equipment consists of 60 CRT terminals which are linked to the University Computing Center's Honeywell DPS-8; 8 color graphics terminals; Tektronix 4051, 4052, and 4081 intelligent graphics systems; two flat bed plotters; a Cromemco System-2 microcomputer; 150 Tandy, Apple, and/or IBM microcomputers; and a DEC PDP-11/34. In addition, the college is currently in the process of acquiring an additional super-minicomputer and has recently purchased 100 additional microcomputers.

The computer engineering technology program is supported by the following equipment: GENRAD microdevelopment system with capability to generate EPROMS for M6800 or Z80 microprocessors, M6800D2 microcomputer, KIM-1 6502 microprocessor, three M6809 microprocessors, three TI 990/189 university board microcomputers, one 16-channel logic analyzer, numerous logic trainers, oscilloscopes and bench meters, and an Intel SDK-2920 analog development system.

4.5 Additional Equipment Required

As noted in Section 2.6.1, the technology transfer workstation is being developed on an IBM compatible microcomputer. One of these units is currently available; however, a second unit is required to meet the schedule of Figure 10. The workstation will consist of a suitably configured 16-bit microcomputer using an MS-DOS operating system. The

microcomputer will include: a minimum 512 kb of main memory, 20 mb fixed disk drive, 360 kb removable disk drive, 1200 baud modem, string comparator coprocessor, and a graphics printer.

5. EXPERIENCE AND PAST PERFORMANCE

The following table describes contracts which involve software development activities carried-out by Dr. Huffman.

TABLE 3
PREVIOUS SOFTWARE DEVELOPMENT CONTRACTS

CONTRACT TITLE	CONTRACTING AGENCY	CONTRACT PERIOD	CONTRACT AMOUNT	BRIEF DESCRIPTION
Calibration and Software Development for an Intermittent Blow-Down Wind Tunnel	Naval Weapons Support Center	June 5, 1981- June 30, 1983	\$ 68,178	Development of real-time control soft- ware for an intermittent blow-down wind tunnel used for decoy flare testing.
Engineering Technical Support	Naval Avionics Center	Dec. 8, 1981- Nov. 30, 1984	\$ 3,321,707	A multi-task program with limited soft- ware activi- ties. Soft- ware tasks consisted of designing and developing control algor- ithms using Ada.
Development of a Heating System, Life Cycle Cost Computer Program	Indiana Gas and Citizens Gas	Aug. 3, 1981- March 17, 1982	\$ 53,292	Development of computer programs for comparison of life-cycle costs for various heat- ing and air conditioning systems.
Computer Data Acquisition System	Ford Motor Company	Nov. 1, 1981- April 16, 1982	\$ 37,047	Development of real-time control and data analysis software for an automotive radiator test facility.

TABLE 3 (continued)
PREVIOUS SOFTWARE DEVELOPMENT CONTRACTS

CONTRACT TITLE	CONTRACTING AGENCY	CONTRACT PERIOD	CONTRACT AMOUNT	BRIEF DESCRIP- TION
Computer Assisted Operating System	VERCO	Jan. 1, 1983- Dec. 30, 1983	\$ 50,000	Development of optimal control algor- ithms for industrial incinerators.
Development of Digital Dynamometer Control Systems	Pontiac Motor Company	Mar. 17, 1982- Feb. 28, 1983	\$ 75,000	Development of computer algor- ithms for the control and data analysis of automotive dynamometers.
Technology Transfer Program	NASA	July 1, 1983- June 30, 1984	\$ 676,000 ⁽¹⁾	Operation of the Aerospace Research Application Center (ARAC).
Technology Transfer Program	NASA	July 1, 1984- June 30, 1985	\$ 1,230,000 ⁽¹⁾	Operation of ARAC and development of software support tools for technology transfer.
Software Development for Technology Transfer	NASA ⁽²⁾	July 1, 1984- June 30, 1985	\$ 86,163	Development of software support tools for technology transfer.

(1) NASA funding and client income.

(2) Subcontracted through the Indianapolis Center for Advanced Research

TABLE 3 (continued)
PREVIOUS SOFTWARE DEVELOPMENT CONTRACTS

Software Development for Technology Transfer	NASA	July 1, 1985- June 30, 1986	\$ 99,925	Development of software support tools for technology transfer
Conversion of SORT-AID for operation on a DEC VAX II-780	Lawrence Livermore National Laboratory	June 1, 1986- Sept. 30, 1986	\$ 18,000	Development of software tools for technology transfer

6. REFERENCES AND BIBLIOGRAPHY

- ANSI/X3/SPARC Database System Study Group, "Reference Models for DBMs Standardization," private communication, October, 1984.
- Bertrand, D., and Bader, C. R., "Storage and Retrieval of Bibliographic References Using a Microprocessor System," Int. J. Bio-Medical Computing, Vol. 11, 1980
- Bieber, H., "Technology Transfer in Practice," IEEE Trans on Engr. Management, Vol. EM-26, 1969, pp. 144-147.
- Bramwell, Bob, "BROWSE: An On-line Manual and System Without an Acronym," Asterisk, December, 1983.
- Burlem, W.S., "Motivations for Technology Transfer," J. Tech Trans, Vol. 1, 1977, pp. 53-61.
- Bush, V., "As We May Think," Atlantic Monthly, 176, July, 1945.
- Carroll, John M., and Carrithers, Caroline, "Training Wheels in a User Interface," Comm. of the ACM, August, 1984.
- Chapman, R., "NASA Partnership with Industry: Enhancing Technology Transfer," Denver Research Institute Report, 1983.
- Chen, Violet, M.S. Project Report, "An Algorithm for Recognizing and Reformatting Substrings," University of Southern Mississippi, Dept. of Computer Science, 1985.
- CODASYL, "A Status Report on the Activities of the CODASYL End User Facilities Committee," SIGMOD Record, Vol. 10, Nos. 2 and 3, 1978.
- Digman, L.A., "Organizational Factors Related to Technology Transfer and Innovation," J. Tech Trans, Vol. 2, 1977, pp. 61-68.
- Elsasser, T. and Schneider, R., "Two Conflicting Approaches to Adaptation of Technology Transfer, Part I," J. Tech Trans, Vol. 2, 1977, pp. 69-75.
- George, P., Jolly, J. A. and Creighton, J. W., "The Linkers Contribution to Technology Transfer," J. Tech. Tran., Vol. 3, 1978, pp. 51-61.
- Goehring, M. F., "Automating the Search Process-Phase I SORT-AID Implementation and Characterization," Aerospace Research Applications Center, Indianapolis, IN, February 1986.

- Grosch, A. N. "Configuring a Professional Microsystem for Information Processing," Microcomputers for Information Management, Vol. 1, 1984, pp. 15-29.
- Hampel, V. E., et al., "Down Loading and Post-Processing of Bibliographic Information with the TIS Intelligent Gateway Computer," On-Line '82 Conference, November 1-3, 1982.
- Heimbigner, Dennis, "Towards an Integrated Environment for Accessing External Databases," Proceedings of SECIBD ACM-SIGOA Conference on Office Information Systems, June 25-27, 1984, Toronto, Canada, Vol. 5 nos. 1-2.
- Herzog, A. J., "The 'Gatekeeper' Hypothesis and the International Transfer of Scientific Knowledge," J. Tech. Trans, Vol. 6, 1981, pp. 57-72.
- Howard, Steven, M. S. Project Report, "A Tool for Experimenting with Using Hypertext for Determining the Relevance of Citations," University of Southern Mississippi, Dept. of Computer Science, 1985.
- Huffman, G. David, et al., "Aerospace Research Applications Center. Volume I. Technical Proposal," Indianapolis Center for Advanced Research, Inc., Indianapolis, IN, June 1983.
- Huffman, G. David, et al., "Development of a Technology Transfer Workstation," Hawaii International Conference on System Science, January, 1986.
- Huffman, G. David, "Technology Transfer--First Quarterly Report," Indianapolis Center for Advanced Research, Inc., Indianapolis, IN, December 1983.
- Huffman, G. David and Ulrich, J. M., "Technology Transfer--Second Quarterly Report," Indianapolis Center for Advanced Research, Inc., Indianapolis, IN, January 1984.
- Huffman, G. David, "Technology Transfer--Third Quarterly Report," Indianapolis Center for Advanced Research, Inc., Indianapolis, IN, April 1984.
- Huffman, G. David, et al., "Aerospace Research Applications Center. Volume I. Technical Proposal," Indianapolis Center for Advanced Research, Inc., Indianapolis, IN, April 1984.
- Huffman, G. David, et al., "The Industrial Applications Study: A Mechanism for Technology Transfer," The Technology Transfer Society, June 25-26, 1984.

- Huffman, G. David, "A Proposal for the Support of the Indianapolis Center for Advanced Research's Technology Transfer Program," University of Southern Mississippi, June 1984.
- Huffman, G. David, et al., "Technology Transfer--Annual Report," Indianapolis Center for Advanced Research, Inc., Indianapolis, IN, July 1984.
- Huffman, G. David, "Technology Transfer Mechanisms," submitted to J. Tech. Trans., January, 1985.
- Huffman, G. David, "Software Support Systems for Use in Technology Transfer," J. Tech. Trans., Volume 9, 1985, pp. 29-41.
- Huffman, G. David, "Economic Development through Technology Transfer--Mississippi's Approach," Annual Technology Transfer Symposium, San Francisco, CA, June, 1985.
- Janis, F. T., "Economic Development Through Technology Transfer, People Interaction--The Key to Technology Transfer," The Technology Transfer Society, June 20-22, 1983.
- Jervis, P., "Innovation and Technology Transfer--The Roles and Characteristics of Individuals," IEEE Trans. on Engr. Management, Vol. EM-22, 1975, pp. 19-27.
- Johnson, F. D., "Technology Transfer--A View of What Works," J. Tech. Trans. Vol. 7, 1983, pp. 1-4.
- Jolly, J. A. and Creighton, J. W., "The Technology Transfer Process: Concepts, Framework and Methodology," J. Tech. Trans. Vol. 1, 1977, pp. 77-91.
- Jung, W., "Barriers to Technology Transfer and Their Elimination," J. Tech. Trans., Vol. 4, 1980, p. 15-25.
- Jung, W., "Basic Concepts for the Evaluation of Technology Transfer Benefits," J. Tech. Trans., Vol. 5, 1980, pp. 37-49.
- Lefkovitz, David, "PRIMATE: A Microcomputer System for Host Mediation and Private File Text Search," Proceedings of the ASIS Annual Meeting, Volume 19, 1982.
- Marcus, Richard S., Kugel, Peter and Benenfeld, Alan R., "Catalog Information and Text as Indicators and Relevance," JASIS, January, 1978.

- Marcus, Richard S., "User Assistance in Bibliographic Retrieval Networks through a Computer Intermediary," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-12, No. 2, March/April 1982.
- Meadow, Charles T., "The Computer as a Search Intermediary," On-Line, July, 1979.
- Menlo Corporation, "IN-SEARCH" An On-line Database Support Package," 1984.
- O'Keefe, T. G., "Evaluations of Technology Transfer: Some Problems and Solutions," J. Tech. Trans., Vol. 6, 1982, pp. 53-57.
- Price, Lynn A., "Thumb: An Interactive Tool for Accessing and Maintaining Text," IEEE Trans. Systems, Man, and Cybernetics, Vol. SMC-12, March/April, 1982.
- Robertson, D., McCracken, D., and Newell, A., "The ZOG approach to man-machine communication," International Journal of Man-Machine Studies, (1981), 14, 461-488.
- Rothwell, R., "Some Problems of Technology Transfer into Industry: Examples from the Textile Machinery Sector," IEEE Trans. on Engr. Management, Vol. EM-25, 1978, pp. 15-20.
- Rouse, Sandra H., and Rouse, William B., "Computer-Based Manuals for Procedural Information," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-10, August, 1980.
- Sagal, M. W., "Effective Technology Transfer--From Laboratory to Production Line," J. Tech. Trans., Vol. 1, 1977, pp. 7-15.
- Salton, G., "A Theory of Indexing," Society for Industrial and Applied Mathematics, Philadelphia, PA, 1975.
- Salton, G., Automatic Information Organization and Retrieval, McGraw-Hill Book Company, 1968.
- Salton, G., Dynamic Information and Library Processing, Prentice-Hall, Inc. 1975.
- Salton, G. and McGill, M. J., "Introduction to Modern Information Retrieval", McGraw-Hill Book Company, 1983.
- Shannon, C. E., "A Mathematical Theory of Communication," Bell Systems Tech. J., Vol. 27, pp. 379-423 and pp. 623-656, 1948.

- Sheahen, T. P. and Ryan, J. F., "Technology Transfer Into Capital-Intensive Industry: The Paper Industry by the Year 2000," J. Tech. Trans., Vol. 7, 1983, pp. 69-84.
- Steinbach, B. Z. and Steinbach, K. H., "Mission and Responsibility of the Technology Transfer Agent," J. Tech. Trans., Vol. 3, 1978, pp. 67-71.
- Strawn, G. W., "Technology Transfer: What is It? What Can It Do for Me?" J. Tech. Trans., Vol. 6, 1982, pp. 45-51.
- Tinsley, E. C., "Computer Aided Sorting in Industrial Applications Studies," Indianapolis Center for Advanced Research, Inc., Indianapolis, IN, October 1983.
- Treu, Siegfried, "Uniformity in User-Computer Interaction Languages: A Compromise Solution," Int. J. Man-Machine Studies, Vol. 16, 1982.
- University of Southern Mississippi, Graduate Bulletin, 1984.
- van Rijsdergen, C. J. Information Retrieval, Butterworths, 1980.
- Vital, Dennis, M. S. Thesis, "A Portable Database Searcher's Software Toolset," University of Southern Mississippi, Dept. of Computer Science, 1985.
- Weyer, Stephen A., "The Design of a Dynamic Book for Information Search," Int. J. Man-Machine Studies, 17, 1982.
- Wilde, Dan, private communication, October, 1984.
- Witten, Ian H., and Bramwell, Bob, "A System for Interactive Viewing of Structured Documents," Comm. of the ACM, March, 1985.